Rec'd PCT/PTO   3 JUN 2005

REPLICATION PROTEIN

This invention relates to a screening method for the identification of agents which modulate the activity of a DNA replication protein as a target for intervention in cancer
5   therapy and includes agents which modulate said activity.

The invention also relates to the use of the DNA replication protein, and its RNA transcripts in the prognosis and diagnosis of proliferative disease e.g., cancer.

10   **Introduction**

Initiation of DNA replication is a major control point in the mammalian cell cycle, and the point of action of many gene products that are mis-regulated in cancer (Hanahan and Weinberg, 2000). The initiation process involves assembly of pre-replication complex proteins, which include the origin recognition complex (ORC), Cdc6, Cdt1 and Mcm
15   proteins, at replication origins during G1 phase of the cell cycle. This is followed by the action of a second group of proteins, which facilitate loading of DNA polymerases and their accessory factors including PCNA, and the transition to S phase. The initiation process is regulated by cyclin-dependent protein kinase 2 (Cdk2), Cdc7-dbf4 and the Cdt1 inhibitor geminin (for review see Bell and Dutta, 2002). In the nucleus of S phase
20   cells, replication forks cluster together to form hundreds of replication 'foci' or factories (Cook, 1999). Replication factories appear to be linked to a structural framework within the nucleus, however the nature of the molecules that form the link and their role in replication fork activity remains unclear.

1

Identification of proteins involved in eukaryotic DNA replication and analysis of the basic pathways that regulate their activity during the cell cycle has been driven largely by yeast genetics. These proteins and pathways are generally conserved from yeast to

5   man. However, in multi-cellular organisms that differentiate down diverse developmental pathways, additional layers of complexity are being uncovered. For example, in vertebrates several proteins involved in neuronal differentiation also regulate the G1-S phase transition (Ohnuma et al., 2001). These include the cdk inhibitor p21$^{CIP1/WAF1/SDI1}$ which has been implicated in oligodendrocyte differentiation following

10   growth arrest (Zezula et al., 2001), and in the terminal differentiation of other cell types (Parker et al., 1995).

Initiation of DNA replication can be reconstituted in vitro with isolated nuclei and cytosolic extracts from mammalian cells (Krude, 2000; Krude et al., 1997; Laman et al.,

15   2001; Stoeber et al., 1998). Furthermore, using recombinant Cdk2 complexed with either cyclins E or A, replication complex assembly and activation of DNA synthesis can be reconstituted independently (Coverley et al., 2002). We have studied the activation step, catalysed in vitro by cyclin A-cdk2, and shown that a relatively unstudied protein, p21-Cip1 interacting zinc-finger protein (Ciz1) functions during this stage of the initiation

20   process. Human Ciz1 was previously identified using a modified yeast two-hybrid screen with cyclin E-p21, and biochemical analysis supported an interaction with p21 (Mitsui et al., 1999). A potential role in transcription was proposed but not demonstrated, and no other function was assigned to Ciz1. More recently the Ciz1 gene

2

was isolated from a human medulloblastoma derived cDNA library using an in vivo tumorigenesis model (Warder and Keherly, 2003). Our analysis shows for the first time that Ciz1 plays a positive role in initiation of DNA replication.

5   A number of changes to chromatin bound proteins occur when DNA synthesis is activated *in vitro* by recombinant cyclin A-cdk2. The present invention relates to the finding that a cdc6-related antigen, p85, correlates with the initiation of DNA replication and is regulated by cyclin A-cdk2. The protein was cloned from a mouse embryo library and identified as mouse Ciz1.

10

In vitro analysis has shown that Ciz1 protein positively regulates initiation of DNA replication and that its activity is modulated by cdk phosphorylation at threonine 191/2, linking it to the cdk-dependent pathways that control initiation. The Embryonic form mouse Ciz1 is alternately spliced, compared to predicted and somatic forms. Human

15   Ciz1 is also alternately spliced, with variability in the same exons as mouse Ciz1. It has been found that recombinant embryonic form Ciz1 promotes initiation of mammalian DNA replication and that pediatric cancers express 'embryonic-like' forms of Ciz1. Without wishing to be held to one theory, the inventors propose that Ciz1 mis-splicing produces embryonic-like forms of Ciz1 at inappropriate times in development. This

20   promotes inappropriately regulated DNA replication and contributes to formation. or progression of cancer cell lineages.

A number of techniques have been developed in recent years which purport to specifically ablate genes and/or gene products. For example, the use of anti-sense nucleic acid molecules to bind to and thereby block or inactivate target mRNA molecules is an effective means to inhibit the production of gene products.

5

A much more recent technique to specifically ablate gene function is through the introduction of double stranded RNA, also referred to as inhibitory RNA (RNAi), into a cell which results in the destruction of mRNA complementary to the sequence included in the RNAi molecule. The RNAi molecule comprises two complementary strands of

10   RNA (a sense strand and an antisense strand) annealed to each other to form a double stranded RNA molecule. The RNAi molecule is typically derived from the exonic or coding sequence of the gene which is to be ablated.

Nucleic acids and proteins have both a linear sequence structure, as defined by their base

15   or amino acid sequence, and also a three dimensional structure which in part is determined by the linear sequence and also the environment in which these molecules are located. Conventional therapeutic molecules are small molecules, for example, peptides, polypeptides, or antibodies, which bind target molecules to produce an agonistic or antagonistic effect. It has become apparent that nucleic acid molecules also

20   have potential with respect to providing agents with the requisite binding properties which may have therapeutic utility. These nucleic acid molecules are typically referred to as aptamers. Aptamers are small, usually stabilised, nucleic acid molecules which comprise a binding domain for a target molecule.

Aptamers may comprise at least one modified nucleotide base. The term "modified nucleotide base" encompasses nucleotides with a covalently modified base and/or sugar. For example, modified nucleotides include nucleotides having sugars which are

5 covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified nucleotides may also include 2' substituted sugars such as 2'-O-methyl-; 2-O-alkyl; 2-O-allyl; 2'-S-alkyl; 2'-S-allyl; 2'- fluoro-; 2'-halo or 2;azido-ribose, carbocyclic sugar analogues a-anomeric sugars; epimeric sugars such as arabinose, xyloses or lyxoses,

10 pyranose sugars, furanose sugars, and sedoheptulose.

Modified nucleotides are known in the art and include by example and not by way of limitation; alkylated purines and/or pyrimidines; acylated purines and/or pyrimidines; or other heterocycles. These classes of pyrimidines and purines are known in the art and

15 include, pseudoisocytosine; N4, N4-ethanocytosine; 8-hydroxy-N6-methyladenine; 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil; 5-fluorouracil; 5-bromouracil; 5-carboxymethylaminomethyl-2-thiouracil; 5-carboxymethylaminomethyl uracil; dihydrouracil; inosine; N6-isopentyl-adenine; 1-methyladenine; 1-methylpseudouracil; 1-methylguanine; 2,2-dimethylguanine; 2-methyladenine; 2-methylguanine; 3-

20 methylcytosine; 5-methylcytosine; N6-methyladenine; 7-methylguanine; 5-methylaminomethyl uracil; 5-methoxy amino methyl-2-thiouracil; β-D-mannosylqueosine; 5-methoxycarbonylmethyluracil; 5-methoxyuracil; 2 methylthio-N6-

isopentenyladenine; uracil-5-oxyacetic acid methyl ester; psueouracil; 2-thiocytosine; 5-methyl-2 thiouracil, 2-thiouracil; 4-thiouracil; 5-methyluracil; N-uracil-5-oxyacetic acid methylester; uracil 5—oxyacetic acid; queosine; 2-thiocytosine; 5-propyluracil; 5-propylcytosine; 5-ethyluracil; 5-ethylcytosine; 5-butyluracil; 5-pentyluracil; 5-pentylcytosine; and 2,6,-diaminopurine; methylpsuedouracil; 1-methylguanine; 1-methylcytosine;

Aptamers may be synthesized using conventional phosphodiester linked nucleotides using standard solid or solution phase synthesis techniques which are known in the art. Linkages between nucleotides may use alternative linking molecules. For example, linking groups of the formula P(O)S, (thioate); P(S)S, (dithioate); P(O)NR'2; P(O)R'; P(O)OR6; CO; or CONR'2 wherein R is H (or a salt) or alkyl (1-12C) and R6 is alkyl (1-9C) is joined to adjacent nucleotides through –O- or –S-.

Other techniques which purport to specifically ablate genes and/or gene products focus on modulating the function or interfering with the activity of protein molecules. Proteins can be targeted by chemical inhibitors drawn, for example, from existing small molecule libraries.

Antibodies, preferably monoclonal, can be raised for example in mice or rats against different protein isoforms. Antibodies, also known as immunoglobulins, are protein molecules which have specificity for foreign molecules (antigens). Immunoglobulins (Ig) are a class of structurally related proteins consisting of two pairs of polypeptide

chains, one pair of light (L) (low molecular weight) chain (κ or λ), and one pair of heavy

(H) chains (γ, α, μ, δ and ε), all four linked together by disulphide bonds. Both H and L

chains have regions that contribute to the binding of antigen and that are highly variable

from one Ig molecule to another. In addition, H and L chains contain regions that are

5   non-variable or constant.

The L chains consist of two domains. The carboxy-terminal domain is essentially

identical among L chains of a given type and is referred to as the "constant" (C) region.

The amino terminal domain varies from one L chain to anther and contributes to the

10   binding site of the antibody. Because of its variability, it is referred to as the "variable"

(V) region.

The H chains of Ig molecules are of several classes, α, μ, σ, α, and γ (of which there are

several sub-classes). An assembled Ig molecule consisting of one or more units of two

15   identical H and L chains, derives its name from the H chain that it possesses. Thus,

there are five Ig isotypes: IgA, IgM, IgD, IgE and IgG (with four sub-classes based on

the differences in the H chains, i.e., IgG1, IgG2, IgG3 and IgG4). Further detail

regarding antibody structure and their various functions can be found in, Using

Antibodies: A laboratory manual, Cold Spring Harbour Laboratory Press.

20

Chimeric antibodies are recombinant antibodies in which all of the V-regions of a mouse

or rat antibody are combined with human antibody C-regions. Humanised antibodies are

recombinant hybrid antibodies which fuse the complimentarity determining regions from a rodent antibody V-region with the framework regions from the human antibody V-regions. The C-regions from the human antibody are also used. The complimentarity determining regions (CDRs) are the regions within the N-terminal domain of both the

5   heavy and light chain of the antibody to where the majority of the variation of the V-region is restricted. These regions form loops at the surface of the antibody molecule. These loops provide the binding surface between the antibody and antigen.

Antibodies from non-human animals provoke an immune response to the foreign

10   antibody and its removal from the circulation. Both chimeric and humanised antibodies have reduced antigenicity when injected to a human subject because there is a reduced amount of rodent (i.e. foreign) antibody within the recombinant hybrid antibody, while the human antibody regions do not illicit an immune response. This results in a weaker immune response and a decrease in the clearance of the antibody. This is clearly

15   desirable when using therapeutic antibodies in the treatment of human diseases. Humanised antibodies are designed to have less "foreign" antibody regions and are therefore thought to be less immunogenic than chimeric antibodies.

Other techniques for targetting at the protein level include the use of randomly generated

20   peptides that specifically bind to proteins, and any other molecules which bind to proteins or protein variants and modify the function thereof.

Understanding the DNA replication process is of prime concern in the field of cancer

therapy. It is known that cancer cells can become resistant to chemotherapeutic agents

and can evade detection by the immune system. There is an on going need to identify

targets for cancer therapy so that new agents can be identified. The DNA replication

5   process represents a prime target for drug intervention in cancer therapy. There is a

need to identify gene products which modulate DNA replication and which contribute to

formation or progression of cancer cell lineages, and to develop agents that affect their

function.

10   **Statements of the invention**

According to one aspect of the present invention there is provided the use of a Ciz1

nucleotide or polypeptide sequence, or any fragment or variant thereof, as a target for

the identification of agents which modulate DNA replication.

15   As used herein the term 'fragment' or 'variant' is used to refer to any nucleic or amino

acid sequence which is derived from the full length nucleotide or amino acid sequence

of Ciz1 or derived from a splice variant thereof. In one embodiment of the invention the

fragment is of sufficient length and/or of sufficient homology to full length Ciz1 to

retain the DNA replication activity of Ciz1. In an alternative embodiment inactive Ciz1

20   fragments are used. The term 'fragment' or 'variant' also relates to the Ciz1 RNA

transcripts described herein and protein isoforms (or parts thereof).

9

As used herein the term 'modulate' is used to refer to either increasing or decreasing

DNA replication, above and below the levels which would normally be observed in the

absence of the specific agent (i.e., any alterations in DNA replication activity which are

either directly or indirectly linked to the use of the agent). The term 'modulate' also

5   includes reference to a change of spacial or temporal organisation of DNA replication.

According to an alternative aspect of the invention there is provided a screening method

for the identification of agents which modulate DNA replication wherein the screening

method comprises the use of Ciz1 nucleotide or polypeptide sequence or fragments or

10   variants thereof.

Preferably the screening method comprises detecting or measuring the effect of an agent

on a nucleic acid molecule selected from the groups consisting of:

        a) a nucleic acid molecule comprising a nucleic acid sequence represented in any

15   of Figures 14, 15, or 21;

        b) a nucleic acid molecule which hybridises to the nucleic acid sequence in (a)

        and which has Ciz1 activity or activity of a variant thereof;

        c) a nucleic acid molecule which has a nucleic acid sequence which is degenerate

        because of the genetic code to the sequences in a) and b) and a candidate agent to

20   be tested;

        d) a nucleic acid molecule derived from the genomic sequence at the Ciz1 locus

        or a nucleic acid molecule that hybridises to the genomic sequence.

In one embodiment of the invention, the nucleic acid molecule is modified by deletion, substitution or addition of at least one nucleic acid residue of the nucleic acid sequence.

Alternatively the screening method comprises the steps of:

5   (i)   forming a preparation comprising a polypeptide molecule, or an active fragment thereof, encoded by a nucleic acid molecule selected from the group consisting of:

a) a nucleic acid molecule comprising a nucleic acid sequence represented in Figs 14, 15 or 21;

b) a nucleic acid molecule which hybridises to the nucleic acid sequence in (a)

10   and which has Ciz1 activity or activity of a variant thereof;

c) a nucleic acid molecule which has a nucleic acid sequence which is degenerate because of the genetic code to the sequences in a) and b) and a candidate agent to be tested;

d) a nucleic acid molecule derived from the genomic sequence at the Ciz1 locus

15   or a nucleic acid molecule that hybridises to the genomic sequence; and

ii)   detecting or measuring the effect of the agent on the activity of said polypeptide.

Assays for the detection of DNA replication are known in the art. Activity residing in Ciz1, or derived peptide fragments, and the effect of potential therapeutic agents on that

20   activity would be assayed *in vitro* or *in vivo*.

In vitro assays for Ciz1 protein activity would comprise synchronised isolated G1 phase nuclei and either S phase extract or G1 phase extract supplemented with cyclin-

11

dependent kinases. Inclusion of Ciz1 or derived peptide fragments stimulates initiation of DNA replication in these circumstances and can be monitored visually (by scoring nuclei that have incorporated fluorescent nucleotides during *in vitro* reactions) or by measuring incorporation of radioactive nucleotides. The assay for therapeutic reagents that interfere with Ciz1 protein function would involve looking for inhibition of DNA replication in these assays. The effect of agents on Ciz1 nuclear localisation, chromatin binding, stability, modification and protein-protein interactions could also be monitored in these assays.

In vivo assays will include creation of cell and mouse models that over-express or under-express Ciz1, or derived fragments, resulting in altered cell proliferation. The preparation of transgenic animals is generally known in the art and within the ambit of the skilled person. The assay for therapeutic reagents would involve analysis of cell-cycle time, initiation of DNA replication and cancer incidence in the presence and absence of drugs that either impinge on Ciz1 protein activity, or interfere with Ciz1 production by targeting Ciz1 and its variants at the RNA level.

In a preferred method of the invention said hybridisation conditions are stringent.

Stringent hybridisation/washing conditions are well known in the art. For example, nucleic acid hybrids that are stable after washing in 0.1xSSC,0.1% SDS at 60°C. It is well known in the art that optimal hybridisation conditions can be calculated if the sequence of the nucleic acid is known. Typically, hybridisation conditions uses 4 – 6 x

SSPE (20x SSPE contains 175.3g NaCl, 88.2g NaH$_2$PO$_4$ H$_2$O and 7.4g EDTA dissolved

to 1 litre and the pH adjusted to 7.4); 5-10x Denhardts solution (50x Denhardts solution

contains 5g Ficoll (Type 400, Pharmacia), 5g polyvinylpyrrolidone abd 5g bovine serum

albumen; 100μg-1.0mg/ml sonicated salmon/herring DNA; 0.1-1.0% sodium dodecyl

5   sulphate; optionally 40-60% deionised formamide. Hybridisation temperature will vary

depending on the GC content of the nucleic acid target sequence but will typically be

between 42^0- 65^0 C.

In a preferred method of the invention said polypeptide is modified by deletion,

10   substitution or addition of at least one amino acid residue of the polypeptide sequence.

A modified or variant, i.e. a fragment polypeptide and reference polypeptide, may differ

in amino acid sequence by one or more substitutions, additions, deletions, truncations

which may be present in any combination. Among preferred variants are those that vary

15   from a reference polypeptide by conservative amino acid substitutions. Such

substitutions are those that substitute a given amino acid by another amino acid of like

characteristics. The following non-limiting list of amino acids are considered

conservative replacements (similar): a) alanine, serine, and threonine; b) glutamic acid

and asparatic acid; c) asparagine and glutamine d) arginine and lysine; e) isoleucine,

20   leucine, methionine and valine and f) phenylalanine, tyrosine and tryptophan. Preferred

are variants which retain the same biological function and activity as the reference

polypeptide from which it varies. Alternatively, variants include those with an altered

13

biological function, for example variants which act as antagonists, so called "dominant negative" variants.

5   Alternatively or in addition, non-conservative substitutions may give the desired biological activity see Cain SA, Williams DM, Harris V, Monk PN. Selection of novel ligands from a whole-molecule randomly mutated C5a library. Protein Eng. 2001 Mar;14(3):189-93, which is incorporated by reference.

A functionally equivalent polypeptide sequence according to the invention is a variant
10   wherein one or more amino acid residues are substituted with conserved or non-conserved amino acid residues, or one in which one or more amino acid residues includes a substituent group. Conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxl residues Ser and Thr; exchange of the acidic residues Asp and Glu; substitution
15   between amide residues Asn and Gln; exchange of the basic residues Lys and Arg; and replacements among aromatic residues Phe and Tyr.

In addition, the invention features nucleotide or polypeptide sequences having at least 50% identity with the nucleotide or polypeptide sequences as hereindisclosed, or fragments and functionally equivalent polypeptides thereof.   In one embodiment, the
20   nucleotide or polypeptide sequences have at least 75% to 85% identity, more preferably at least 90% identity, even more preferably at least 95% identity, still more preferably at

least 97% identity, and most preferably at least 99% identity with the nucleotide and amino acid sequences illustrated herein.

In a preferred method of the invention said nucleic acid molecule comprises the nucleic acid sequence encoding the amino acid sequence Ciz1 in Fig 16 or Fig 17 or any variants thereof, including those described in Figures 20A and 20B . In a further preferred method of the invention said nucleic acid molecule consists of the nucleic acid sequence which encodes the amino acid sequence Ciz1 in Fig 16 or Fig 17 or variants thereof, including those described in Figures 20A and 20B.

In a further preferred method of the invention said polypeptide molecule comprises the amino acid sequence Ciz1 in Fig 16 or 17 or variants thereof, including those described in Figures 20A and 20B . In a further preferred method of the invention said polypeptide molecule consists of the amino acid sequence Ciz1 in Fig 16 or 17 or variants thereof, including those described in Figures 20A and 20B.

In a further preferred method of the invention said polypeptide is expressed by a cell, preferably a mammalian cell, or animal and said screening method is a cell-based screening method.

Preferably said cell naturally expresses the Ciz 1 polypeptide. Alternatively said cell is transfected with a nucleic acid molecule encoding a Ciz 1 polypeptide (or a variant molecule thereof, found, for example in cancer cell lineages).

According to a further aspect of the invention there is provided an agent obtainable by the method according to the invention.

5   Preferably said agent is an antagonist of Ciz1 mediated DNA replication. Alternatively said agent is an agonist of Ciz1 mediated DNA replication.

In a further preferred method of the invention said agent is selected from the group consisting of: polypeptide; peptide; aptamer; chemical; antibody; nucleic acid; or

10   polypeptide or nucleotide probe.

Preferably the agent comprises a sequence that is complimentary or of sufficient homology to give specific binding to the target and can be used to detect the level of nucleic acid or protein for diagnostic purposes.

15

Alternatively the agent identified by the method of the invention is a therapeutic agent and can be used for the treatment of disease.

In one embodiment of the invention the agent is an antibody molecule and binds to any

20   of the sequences represented by figures 16, 17 or 20.

Preferably said antibody is a monoclonal antibody.

Alternatively said agent is an anti-sense nucleic acid molecule which binds to and thereby blocks or inactivates the mRNA encoded by any of the nucleic acid sequences in (i) above.

5   In an alternative embodiment, said agent is an RNAi molecule and comprises two complementary strands of RNA (a sense strand and an antisense strand) annealed to each other to form a double stranded RNA molecule. Preferably the RNAi molecule is derived from the exonic sequence of the Ciz1 gene or from another over-lapping gene.

10   In one embodiment unspliced mRNA is targetted with RNAi to inhibit production of the spliced variant. In another the spliced variant mRNA is ablated without affecting the non-variant mRNA.

In a preferred method of the invention said peptide is an oligopeptide. Preferably, said
15   oligopeptide is at least 10 amino acids long. Preferably said oligopeptide is at least 20, 30, 40, 50 amino acids in length.

In a further preferred method of the invention said peptide is a modified peptide.

20   It will be apparent to one skilled in the art that modified amino acids include, by way of example and not by way of limitation, 4-hydroxyproline, 5-hydroxylysine, N^6-acetyllysine, N^6-methyllysine, N^6,N^6-dimethyllysine, N^6,N^6,N^6-trimethyllysine, cyclohexyalanine, D-amino acids, ornithine. Other modifications include amino acids

with a C_2, C_3 or C_4 alkyl R group optionally substituted by 1, 2 or 3 substituents selected from halo (eg F, Br, I), hydroxy or C_1-C_4 alkoxy.

Alternatively said peptide is modified by acetylation and/or amidation.

5

In a preferred method of the invention the polypeptides or peptides are modified by cyclisation. Cyclisation is known in the art, (see Scott *et al* Chem Biol (2001), 8:801-815; Gellerman et al J. Peptide Res (2001), 57: 277-291; Dutta *et al* J. Peptide Res (2000), 8: 398-412; Ngoka and Gross J Amer Soc Mass Spec (1999), 10:360-363).

10

According to a further aspect of the invention there is provided a vector as a delivery means for, for example, an antisense or an RNAi molecule which inhibits Ciz1 or variants thereof and thereby allows the targetting of cells expressing the protein to be targeted.

15

In one embodiment of the invention a viral vector is used as delivery means.

Preferably the vector includes an expression cassette comprising the nucleotide sequence selected from the group consisting of;

20   a) the nucleic acid sequence which encodes Ciz1 amino acid sequence as shown
            in Fig 14, 15 and 21;

            b) a nucleic acid molecule which hybridizes to the nucleic acid sequence of (a) ;

18

c) a nucleic acid molecule which has a nucleic acid sequence which is degenerate because of the genetic code to the sequences in a) and b) and any sequence which is complimentary to any of the above sequences;

d) a nucleic acid sequence that encodes Ciz1 pre-mRNA (i.e., the genomic

5   sequence),

wherein the expression cassette is transcriptionally linked to a promoter sequence.

Preferably the vectors including the expression cassette is adapted for eukaryotic gene

10   expression. Typically said adaptation includes, by example and not by way of limitation, the provision of transcription control sequences (promoter sequences) which mediate cell/tissue specific expression. These promoter sequences may be cell/tissue specific, inducible or constitutive.

15   Promoter elements typically also include so called TATA box and RNA polymerase initiation selection sequences which function to select a site of transcription initiation. These sequences also bind polypeptides which function, *inter alia*, to facilitate transcription initiation selection by RNA polymerase.

20   Adaptations also include the provision of selectable markers and autonomous replication sequences which both facilitate the maintenance of said vector in either the eukaryotic cell or prokaryotic host. Vectors which are maintained autonomously are referred to as

episomal vectors. Further adaptations which facilitate the expression of vector encoded genes include the provision of transcription termination sequences.

These adaptations are well known in the art. There is a significant amount of published literature with respect to expression vector construction and recombinant DNA techniques in general. Please see, Sambrook et al (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY and references therein; Marston, F (1987) DNA Cloning Techniques: A Practical Approach Vol III IRL Press, Oxford UK; DNA Cloning: F M Ausubel et al, Current Protocols in Molecular Biology, John Wiley & Sons, Inc.(1994).

According to the present invention there is provided a diagnostic method for the identification of proliferative disorders comprising detecting the presence or expression of the Ciz1 gene, Ciz1 splice variants and mutations in the genomic or protein sequence thererof.

Preferably said diagnostic method comprises one of more of the following steps:

(i)   contacting a sample isolated from a subject to be tested with an agent which specifically binds a polypeptide with Ciz 1 activity or a nucleic acid molecule encoding a polypeptide with Ciz 1 activity; and

(ii)   detecting or measuring the binding of the agent on said polypeptide or nucleic acid in said sample;

(iii)   use of reverse-transcribed PCR or real-time PCR to monitor Ciz1 isoform

expression and to measure expression levels.

(iv)   measuring the presence of nucleic acid or amino-acid mutations based on altered

conformational properties of the molecule.

5

In one embodiment, the diagnostic method of the present invention is carried out *in-vivo*.

In an alternative embodiment, the diagnostic method of the present invention is carried

out *ex-vivo or in-vitro*.

10

Preferably the diagnostic method provides for a quantitative measure of Ciz1 RNA or

protein variants in a sample.

In one embodiment of the invention there is provided the use of an agent which

15   modulates Ciz1 RNA or protein, or variants thereof, as a pharmaceutical.

Preferably said pharmaceutical comprises an agent identified by the screening method of

the present invention in combination or association with a pharmaceutically acceptable

carrier, excipient or diluent.

20

Preferably said pharmaceutical is for oral or topical administration or for administration

by injection.   In alternative embodiment of the invention the pharmaceutical is

administered as an aerosol.

In a further preferred embodiment of the invention there is provided the use of an agent according to the invention for the manufacture of a medicament for use in the treatment of proliferative disease. Preferably said proliferative disease is cancer.

5

Preferably said cancer is a paediatric cancer and is selected from the group consisting of; retinoblastoma, neuroblastoma, Burkitt lymphoma, medulloblastoma, Ewings Sarcoma family tumours (ESFTs).

10   In an alternative embodiment the cancer is a carcinoma, adenocarcinoma, lymphoma or leukemia.

In an alternate embodiment the disease is liver, lung or skin cancer or metastasis.

15   According to a further aspect of the invention there is provided a method to treat a proliferative disease comprising administering to an animal, preferably a human, an agent obtainable by the method according to the invention.

According to an alternate aspect of the invention, there is provided the use of an agent
20   according to the invention for the manufacture of a medicament to slow cell division or growth.

The invention also includes the use of the Ciz1 amino acid sequence and protein structure in rational drug design and the use of Ciz1 nucleotide and amino acid sequences thereof or variants thereof for screening chemical libraries for agents that specifically bind to Ciz1.

5

The invention also includes a kit comprising a diagnostic, prognostic or therapeutic agent identified by the method of the invention.

In an alternative embodiment of the invention, an array based sequencing chip is used

10   for the detection of altered Ciz1.

Detailed Description of the invention

An embodiment of the invention will now be described by example only and with reference to the following figures:

15

Fig. 1 Illustrates the effect of cyclin A-cdk2 on late G1 nuclei. A) Anti-Cdc6 antibody V1 detects mouse Cdc6 and a second antigen in western blots of 3T3 whole cell extract, which migrates with approximate Mr of 100kDa (based on the mobility of the Mcm3 protein this was previously estimated at nearer 85kDa so the antigen was named p85 –

20   we have kept the same name here for clarity). P85 is present in both the soluble fraction and insoluble nuclear fraction (prepared under in vitro replication conditions). B) Initiation of DNA synthesis in 'replication competent' late G1 phase nuclei by G1 phase extract supplemented with recombinant cyclin A-cdk2. Control bar shows the proportion

of nuclei already in S phase (unshaded), and those that initiated replication in extract

from S phase cells (shaded). C) After 15 minutes under cell-free replication conditions

nuclei were washed and the chromatin fraction was re-isolated and separated by SDS-

Page and blotted for Mcm2 and Mcm3. D) The same nuclei blotted with antibody V1.

5   p85 antigen is more abundant in nuclei exposed to initiation-inducing concentrations of

cyclin A-cdk2. Antibody V1 was used to clone the gene for p85 from a mouse embryo

expression library which was identified as Ciz1.

Fig. 2 Alignment of mouse Ciz1 variants. The predicted full-length Ciz1 amino-acid

10   sequence ('Full') is identical to a mouse mammary tumour cDNA clone (BC018483),

while embryonic Ciz1 ('ECiz1', AJ575057), and a melanoma-derived clone

(AK089986) lack two discrete internal sequences. In addition, the first available

methionine in ECiz1 is in the middle of exon 3 (Met84),which excludes a polyglutamine

rich region from the N-terminus. Melanoma derived AK089986 may be incomplete as it

15   ends 77 codons before the C-terminus of all other mouse and human clones. Stars

indicate amino-acids changed by site-directed mutagenesis in the constructs shown in D.

Amino-acids that correspond to codons targeted by siRNAs are underlined. B) Mouse

Ciz1 is encoded by at least 17 exons. Coding exons are shown in grey, alternatively

spliced regions are black, untranslated regions are white. Two alternative exon 1

20   sequences are included in some Ciz1 transcripts (not shown) but an alternative

translational start site upstream of the two depicted here has not yet been found. C)

Sequence features and putative domains in ECiz1. Predicted nuclear localisation

sequence (NLS), putative cyclin-dependent kinase phosphorylation sites, C2H2 type

zinc-fingers and a C terminal domain with homology to the nuclear matrix protein matrin 3 (Nakayasu and Berezney, 1991) are shown. The positions of sequences absent from ECiz1 are indicated by triangles. D) ECiz1 and derived truncations and point mutants used in cell-free DNA replication experiments. Numbers in parentheses relate to

5   amino-acid positions in the full-length form of mouse Ciz1, shown in A. Stars indicate putative phosphorylation sites ablated by site-directed mutagenesis.

Fig. 3 Shows the effect of Ciz1 protein and derived fragments in cell-free DNA replication experiments and illustrates that ECiz1 promotes initiation of mammalian

10   DNA replication A) Recombinant ECiz1 stimulates initiation of DNA replication in 'replication competent' late G1 phase nuclei, during incubation in S phase extract. Histogram shows the average number of nuclei that incorporated biotinylated nucleotides in vitro (black), in the presence or absence of ectopic ECiz1, with standard deviations calculated from four independent experiments. The17% of nuclei that were

15   already in S phase when the nuclear preparation was made are shown in white. Images show nuclei replicating in vitro, with or without 1 nM ECiz1. Total nuclei are counterstained with propidium iodode (red). B) The response to recombinant ECiz1 is concentration dependent with a sharp optimum in the nM range. In this experiment, and all those shown in B-I, results are expressed as % initiation rather than % replication.

20   This is calculated from the number of nuclei that initiate *in vitro* and the number of nuclei that are 'competent' to initiate *in vitro* (see methods). C) Threonines 191/2 are involved in regulating Ciz1 DNA replication activity as ECiz1 cdk site mutant T(191/2)A escapes suppression at high concentrations. D) Cdk site mutant T(293)A

stimulates initiation with a similar profile to ECiz1 but at lower concentrations. E) Truncated ECiz1 (Nterm 442) lacks C-terminal sequences, but stimulates in vitro initiation to a similar extent as ECiz1. F) Cterm 274 retains no DNA replication activity in this assay. G, H, I) Further deletion analysis in the N-terminal two thirds of the ECiz1

5   protein show that a short region 3' of exon 8 is required for Ciz1 function when assayed in vitro.

Fig. 4 Characterisation of anti-Ciz1 polyclonal antibodies and identification of 125kDa Ciz1-related bands A) Coomassie stained SDS-polyacrylamide gel showing purified

10   recombinant ECiz1 fragment Nterm442, and western blots of recombinant Nterm442 using anti-Cdc6 antibody V1, and anti-Ciz1 antibodies 1793 and 1794. B) Western blot of 3T3 whole cell extract. Of the two bands detected by anti-Ciz1 antibody 1793 one has the same mobility as p85-Ciz1 (100kDa) recognized by antibody V1 and the other has an apparent Mr of 125kDa. Anti-Ciz1 antibody 1794 recognizes only the125kDa form of

15   Ciz1 (and a second antigen of around 80 kDa). C) Immuno-precipitation from 3T3 nuclear extract, using antibody V1 or anti-Ciz1 1793. Both antibodies precipitate p85, which is recognized by the reciprocal antibody in western blots. P125 is precipitated by antibody 1793, and to a lesser extent by antibody V1 and these are recognized by 1793 in western blots. Mcm3 is shown as a control.

20

Fig. 5 Immunofluorescence analysis of endogenous Ciz1. Ciz1 resides in sub-nuclear foci that overlap with sites of DNA replication A) Endogenous Ciz1 (red) in 3T3 cells fixed before (untreated) or after (detergent treated) exposure to TritonX100, detected

with anti-Ciz1 antibody 1793. Nuclei are counterstained with Hoescht 33258 (blue).

Cdc6 (green), detected with a Cdc6-specific monoclonal antibody is shown for

comparison. B) Inclusion of recombinant Ciz1 blocks reactivity of antibody 1793 with

detergent treated nuclei. C) Detergent-resistant Ciz1 (red) is present in all nuclei in

5   cycling populations, while detergent resistant PCNA (green) persists only in S phase

nuclei. D) High magnification confocal sections of detergent resistant Ciz1 and PCNA,

and merged image showing co-localising foci (yellow). E) Line plots of red and green

fluorescence across the merged image in D, at the positions indicated (i and ii). F) Cross-

correlation plot (Rubbi and Milner, 2000; van Steensel et al., 1996) for green foci

10   compared to red over the whole merged image in D, and (inset) for the marked section

after thresh-holding fluorescence at the levels shown in Eii. The red line in the inset to F

shows loss of correlation when the Ciz1 image is rotated $90°$ with respect to PCNA. Bar

is 10 µM.

15   Fig. 6 RNA interference. Ciz1 depletion inhibits S phase A) siRNAs that target Ciz1

transcripts at four sites (see Fig. 2A) were individually applied to cycling 3T3 cells as a

single 3nM dose and cell number was monitored at the indicated times. Images of cell

populations at 16 and 40 hours after transfection with siRNA 8 (red outline) or mock

treated cells (blue outline) are shown. B) Ciz1 protein detected with anti-Ciz1 1793

20   (green) 48 hours after exposure to Ciz1 siRNAs (4 and 8), or control GAPDH siRNA. C)

Ciz1, GAPDH and β-actin transcript levels in cells exposed to Ciz1 siRNAs (4 and 8), or

control GAPDH siRNA for 24 hours. Numbers in parentheses reflect band intensity in

arbitrary units, and the overall reduction in Ciz1 and GAPDH transcripts (normalised

against β-actin) is expressed as a percentage. D) The proportion of cells that

incorporated BrdU into DNA (green) is significantly decreased in Ciz1 depleted cells, 48

hours after treatment with Ciz1 siRNA. Histogram shows average results from four

5   independent experiments. E) The number of nuclei with detergent resistant Mcm3

(green) increases in populations treated with Ciz1 siRNA. F) The proportion of nuclei

with detergent resistant PCNA (green) also increases under these conditions. All nuclei

are counterstained and shown in pseudo-colour (red).

10   Fig. 7 RT-PCR analysis of Ciz1 exons 3/4 splice variant expression in mouse primordial

germ cells and embryonic stem cells. Exons 3 and/or 4 are alternatively spliced in these

cell types, but not in neonatal heart. These data are consistent with the hypothesis that

full-length Ciz1 is the pre-dominant form in neonatal somatic tissue, and that variants

occur with more frequency earlier in development, and in germ line tissues.

15

Fig. 8 Transient transfection of mouse 3T3 cells. A. GFP-tagged Ciz1 constructs were

transfected into NIH3T3 cells or B. microinjected into the male pro-nucleus of fertilized

mouse eggs at the one cell stage. By 24 hours Ciz1 and ECiz1 became localized to the

nucleus forming a subnuclear spotty pattern, while GFP alone was present in both the

20   nucleus and the cytoplasm. C. High magnification images of live 3T3 cell nuclei 24

hours after transfection showing the subnuclear organisation of EGFP tagged Ciz1 and

ECiz1 and derived fragments with the C-terminal fragment (equivelent to Cterm274)

removed. In the absence of C-terminal domains GFP-ECiz1 is diffusely localised in the

nucleus 24 hours after transfection, while GFP-Ciz1 aggregates to form one or two large

blobs within the nucleus. D. The Cterminal 274 domain alone is cytoplasmic until after

cells have passed through mitosis (most likely due to lack of nuclear localisation

5   sequences and passive entry to the nucleus), but once inside binds to nuclear structures

and condenses with chromosomes. E. Representative images of GFP-Ciz1 (green), BrdU

(red) and total nuclei (blue) in a population labelled with BrdU for the first 12 hours

after transfection are shown. Histograms show the proportion of transfected (green) cells

that incorporated BrdU compared to the number of untransfected (grey) cells for three

10   separate labelling windows. During 0-22 hours after transfection rapidly cycling cells

registered a consistent increase in the BrdU labelled fraction when transfected with

either Ciz1 or ECiz1. Similar results were obtained with dense cultures in which most

cells had exited the cell cycle and entered quiescence. However, when rapidly cycling

cells were exposed to BrDu for a short (20 minute) pulse 22 hours after transfection the

15   number of cells engaged in DNA synthesis was reduced in the Ciz1 and ECiz1

transfected populations, compared to untransfected controls and cells transfected with

GFP alone. This indicates that by 22 hours DNA synthesis had ceased in Ciz1

expressing cells.

20   Fig. 9 Altered proliferation potential and cell morphology in transfected populations.

Cell clusters arising in transfected 3T3 cell populations. A. Cells were transfected with

the N-terminal two thirds of Ciz1 or ECiz1 (N-term442) tagged with GFP, and

maintained under selection with 50 µg/ml G418. After three weeks under selection, cell

aggregates were visible with GFP positive cells within.

5 Fig. 10 Human Ciz1 splice variants in paediatric cancers. There are seven human Ciz1

cDNAs in public databases, but only one is derived from normal adult tissue (B cells)

and it contains all predicted exons. The other six are derived from embryonic cells or

paediatric cancers. Five of these are alternatively spliced with variability in exons 2, 3,

6, and 8 (like mouse ECiz1), and also in exon 4 (like mouse ES cells, primordial germ

cells and testis). The sixth (AF159025) lacks the first methionine and contains single-

10 nucleotide polymorphisms that give rise to amino-acid substitutions. All differences

from the predicted sequence (AB030835) are marked.

Fig. 11 EST sequence analysis. On each map a schematic representation of the Ciz1

protein is included for reference, showing the positions of alternatively spliced exons

15 (black), putative chromatin interaction domains (grey) and predicted zinc fingers (black

vertical lines). All EST sequences are accompanied by their Genbank accession number

with the library from which they were derived indicted in parentheses. Sequences absent

from Ciz1 ESTs due to alternative splicing are shown in yellow, frame-shifts in red and

putative deletions in grey. Single nucleotide polymorphisms that give rise to amino-acid

20 substitutions are indicated by black dots and some of these occur in a consensus cdk

phosphorylation site which we have shown to be important for the regulation of Ciz1

30

activity (blue dots). Position of the inserted sequence in the carcinoma cell line MGC102

is indicated by a triangle.

A) Translated ESTs from paediatric cancers and adult neural cancers.

5   B) Translated ESTs from various non-cancer cells and tissues

C) Translated ESTs from leukemias, lymphomas, and from normal haematopoetic

and lymphocytic cells

D) Translated ESTs from carcinomas

E) Translated ESTs from a range of other cancers

10   F) Summary of alternatively spliced regions in human Ciz1 showing conditionally

included sequences.

Fig. 12 Ciz1 splice variant expression in Ewings sarcoma family tumour cells lines

(ESFT) and neuroblastoma cell lines. A. Whole RNA samples from six independent

15   ESFT cell lines, two neuroblastomas and a control cell line (HEK293 cells) was subject

to RT-PCR analysis using 4 different primer sets.

ESFT cell lines are 1) A673, 2) RDES, 3) SKES1, 4) SKNMC, 5) TC3, 6) TTC466.

Neuroblastoma cell lines are 1) IMR32, 2) SKNSH.

B. Analysis of Ciz1 Exons 3/4/5 PCR products in ESFTs and neuroblastoma. The

20   products of primers h3 and h4 (spanning potentially variable exons 4 and 6) were

analysed in more detail. PCR fragments were purified from agarose gels by standard

procedures, subcloned and sequenced to identify the source of fragment size variations.

Between one and eleven individual clones for each of the seven cell lines were

sequenced and the results are summarised in tabular form. Ciz1 from ESFT cell lines

lacks exon 4 in 31% of transcripts overall, and for some ESFT lines this is nearer 50%.

DSSSQ is more commonly absent in the two neuroblastoma cell lines tested here.

5   Fig. 13 Ciz1 isoforms in normal human fibroblasts (Wi38) and metastatic prostate

cancer cell lines (PC3 and LNCAP). A. Both prostate cancer cell lines contain an excess

of the largest p125 Ciz1 protein variant in the nuclear fraction, compared to the non-

cancer cell line. B. Models for the production of p85 (100) from p125 variants by protein

processing during initiation of DNA replication.

10

Fig. 14 illustrates the full length mouse mRNA sequence.

Fig. 15 illustrates the full length human mRNA sequence.

15   Fig. 16 illustrates the full length mouse protein sequence.

Fig. 17 illustrates the full length human protein sequence.

Fig. 18 illustrates human alternatively spliced protein sequences. Sequences shown are

20   absent in the spliced protein sequences.

Fig. 19 illustrates human alternatively spliced mRNA sequences. Sequences shown are

absent in the spliced protein sequences.

Fig. 20 A and B illustrate unique junction sequences created in human Ciz1 proteins by missing exons. Junction sequences represent prime sites of target for therapeutic agents identified by the method of the invention.

5   Fig 21 A to H illustrate junction sequences created in human Ciz1 mRNA.

Identification of Ciz1 We have exploited a polyclonal antibody (antibody V1) that was raised against recombinant human Cdc6 (Coverley et al., 2000; Stoeber et al., 1998;

10   Williams et al., 1998) to identify and study an unknown antigen whose behaviour correlates with initiation of DNA replication in vitro. The antigen has an apparent Mr of 100kDa (called p85) and is readily detectable in extracts from 3T3 cells (Fig. 1A).

DNA synthesis can be activated in cell-free replication experiments using 'replication

15   competent' late G1 phase nuclei, G1 extracts, and recombinant cyclin A-cdk2. Under these conditions nuclei will incorporate labelled nucleotides into nascent DNA, in a manner strictly dependent on the concentration of active protein kinase (Fig. 1B). Above and below the optimum concentration no initiation of DNA replication takes place. However, other events occur which inversely correlate with initiation (Coverley et al.,

20   2002). Here we use activation of DNA synthesis (Fig. 1B), and Mcm2 phosphorylation (which results in increased mobility, Fig. 1C), to calibrate the effects of recombinant cyclin A-cdk2 in cell-free replication experiments, and correlate the behaviour of p85 with activation of DNA synthesis.

33

In G1 nuclei that are re-isolated from reactions containing initiation-inducing concentrations of cyclin A-cdk2, p85 antigen is more prevalent compared to nuclei exposed to lower or higher concentrations of kinase (Fig. 1D). This suggests that p85 is

5   regulated at some level by cyclin A-cdk2, in a manner that is co-incident with activation of DNA synthesis. No other antigens correlate so closely with this stage in the cell-free initiation process, therefore we used antibody V1 to clone the gene for mouse p85.

When applied to a cDNA expression library derived from 11-day mouse embryos

10   antibody V1 picked out two clones that survived multiple rounds of screening (see methods). One encoded mouse Cdc6, while the other encoded 716 amino acids of the murine homologue of human Ciz1 (Mitsui et al., 1999). Full-length human and mouse Ciz1 have approximately 70% overall homology at the amino-acid level, with greatest (>80%) homology in the N and C terminal regions. Ciz1 is conserved among vertebrates

15   as homologues exist in rat and fugu, but no proteins with a high degree of homology or similar domain structure could be identified in lower eukaryotes, raising the possibility that Ciz1 evolved to perform a specialised role in vertebrate development.

A previous publication on human Ciz1 (Mitsui et al 1999) demonstrated interaction with

20   the cell-cycle protein p21-CIP1, leading to investigation of a proposed role as a transcription factor, not a DNA replication factor. A second paper (Warder and Keherly 2003) published after the priority date of this patent application suggests a role for Ciz1

in tumorigenesis, but does not demonstrate a role in DNA replication or recognise the importance of Ciz1 splice variant expression.

Multple Ciz1 isoforms The predicted mouse Ciz1 open reading frame and a cDNA
5   derived from a mouse mammary tumour library (BC018483) contain three regions that are not present in our embryonic clone (AJ575057), hereafter referred to as ECiz1 (Fig. 2A). The three variable regions in ECiz1 appear to be the result of alternative splicing of exons 2/3, 6 and 8 (Fig. 2B). Mouse melanoma clone AK089986 lacks two of the same three regions as ECiz1 (Fig. 2A), while the third encodes an N-terminal polyglutamine
10   stretch that is also absent from human medulloblastoma derived clones. A fourth sequence block derived from exons 3/4 is absent from Ciz1 transcripts derived from mouse ES cells, and from exon 4 in mouse primordial germ cells (fig. 7). Human Ciz1 is also alternatively spliced at the RNA level to yield transcripts that exclude combinations of the same four sequence blocks as mouse Ciz1 (see below). In fact, all known
15   variations in mouse Ciz1 cDNAs have close human parallels, some of which are identical at the amino-acid level. This suggests that the different Ciz1 isoforms have functional significance. A fifth variable regions (not yet observed in the mouse) is alternatively spliced in human Ciz1 transcripts derived mainly from carcinomas.

20   The data suggest that shorter forms of Ciz1 (lacking the alternatively spliced exons) are most prevalent early in development and in cell lineages that give rise to the germ line. In the analysis shown in fig. 7, only Ciz1 from fully developed neonatal heart shows no alternative splicing, while all embryonic cell types contain alternatively spliced forms.

Furthermore, the only complete Ciz1 cDNAs in public databases (human or mouse) are derived from non-embryonic cell types, and the only ones derived from embryonic sources are alternatively spliced. Therefore, Ciz1 splice variant expression appears to occur preferentially in cell types that are not yet fully differentiated.

5

Notably, Ciz1 cDNAs from paediatric cancers are also alternatively spliced (see below). This lead us to the hypothesis that failure to express the appropriate Ciz1 isoform at the right point in development leads to inappropriately regulated Ciz1 activity. This could contribute to unscheduled proliferation and cellular transformation.

10

ECiz1 stimulates DNA replication in vitro Upon exposure to cytosolic extract from S phase cells, late G1 phase nuclei initiate DNA replication and begin synthesizing nascent DNA (Krude et al., 1997). We used this cell-free assay to test the effect of ECiz1, and derived recombinant fragments, on DNA synthesis (Fig. 3). Full-length ECiz1 protein

15   consistently increased the number of nuclei that replicated in vitro, from 30%(+/-0.9%) to 46% (+/- 5.5%), which suggests that Ciz1 is limiting for initiation in S phase extracts (Fig. 3A). Only two other classes of protein (cyclin-dependent kinases , Coverley et al., 2002; Krude et al., 1997; Laman et al., 2001, and the Cdc6 protein, Coverley et al., 2002; Stoeber et al., 1998) have been previously found to stimulate cell-free initiation.

20   Thus, ECiz1 is the first protein to have this property that was not already known to be involved in the replication process. The positive effect of recombinant ECiz1 on cell-free initiation argues that endogenous Ciz1 plays a positive role in DNA replication in mammalian cells.

Stimulation of cell-free initiation is concentration-dependent with peak activity in S

phase extract at around 1nM ECiz1 (Fig. 3B). This echoes previous cell-free analyses

with other recombinant proteins (Coverley et al., 2002; Krude et al., 1997), where

5   stimulation of initiation typically peaks and then falls back to the un-stimulated level at

high concentrations. For ECiz1, the reason for the drop in activity at high concentrations

is not yet clear. However, mutagenesis studies (see below) suggest that the restraining

mechanism is likely to be active and specific rather than due to a general imbalance in

the composition of higher order protein complexes.

10

Down regulation of ECiz1 involves threonines 191/192   Ciz1 is likely to be a phospho-

protein in vivo since it contains numerous putative phosphorylation sites, and it displays

altered mobility when 3T3 cell extracts are treated with lambda phosphatase (not

shown). Murine Ciz1 contains two RXL cyclin binding motifs and five putative cdk-

15   phosphorylation sites, which are present in all known variants. Four of these are located

in the N-terminal fragment of ECiz1 that contains in vitro replication activity (see

below), and one is adjacent to the site at which exon 6 is alternatively spliced to exclude

a short DSSSQ sequence motif (Fig. 2A, C). As this motif is 100% identical and

alternatively spliced in both mouse and man we reasoned that conditional inclusion

20   might serve to regulate Ciz1 activity, identifying this region of the protein as potentially

important. We therefore chose to focus on the cdk site that is four residues upstream and

which is also conserved in mouse and man, by combining a genetic approach with cell-

free replication assays. Starting with ECiz1, two threonines at 191 and 192 were

changed to two alanines, generating ECiz1T(191/2)A (Fig. 2D). When tested in vitro for

DNA replication activity, ECiz1T(191/2)A stimulated initiation in late G1 nuclei to a

similar extent as ECiz1 (Fig. 3C). However unlike ECiz1, stimulation of initiation was

maintained over a broad range of concentrations that extended over at least three orders

5   of magnitude. Therefore, a mechanism to restrict the activity of excess ECiz1 exists and

operates in a cell-free environment. In a separate construct, the threonine at position 293

was also changed to alanine generating ECiz1T(293)A (Fig. 2D), but this alteration had

little effect on ECiz1 activity assayed in vitro (Fig. 3D).

10   These results demonstrate that down-regulation of ECiz1 activity involves threonine

191/2, and is probably caused by cyclin-dependent kinase mediated phosphorylation at

this site. This links Ciz1 activity to the cdk-dependent pathways that control all major

cell-cycle events, including initiation of DNA replication.

15   Most pre-replication complex proteins and many replication fork proteins are

phosphorylated in vivo, often by cyclin-dependent kinases (Bell and Dutta, 2002; Fujita,

1999). Our data suggests that nuclear accumulation of p85-Ciz1 antigen is regulated

(directly or indirectly) by cyclin A-cdk2, and it shows that a specific consensus cdk

phosphorylation site at threonine 191/192 is involved in controlling Ciz1 activity. When

20   this site is made unphosphorylatable Ciz1 activity is maintained over a broader range of

concentrations in cell-free assays. Therefore, Ciz1 activity is normally down regulated

by modification at this site. The functions of the other conserved cdk phosphorylation

sites, and the effect of conditional inclusion of an RXL cyclin-binding motif in the

alternatively spliced N-terminal portion of Ciz1, remain to be determined. Thus, the simple negative relationship between Ciz1 activity and cdk-dependent phosphorylation that has been uncovered here, is unlikely to be the whole story. However, our analysis so far links Ciz1 with the cdk-dependent pathways that control all major cell-cycle

5   transitions, and is therefore consistent with our main conclusion that Ciz1 is involved in initiation of DNA replication.

In vitro replication activity resides in the N-terminus   Ciz1 possesses several C-terminal features that may anchor the protein within the nucleus. The matrin 3 domain suggests

10   interaction with the nuclear matrix and the three zinc-fingers imply interaction with nucleic acids. Indeed, recent evidence suggests that human Ciz1 binds DNA in a weakly sequence specific manner (Warder and Keherley, 2003). To determine whether C-terminal domains are important for ECiz1 replication activity we divided the protein into two fragments (Fig. 2D). Nterm442 (which contains the NLS, two conserved cdk sites,

15   one zinc finger and all known sites where variable splicing has been observed) stimulates initiation to a similar extent and at the same concentration as ECiz1 (Fig. 3E). In contrast, the C-terminal portion (Cterm274) contains no residual replication activity (Fig. 3F). Therefore, the matrin 3 domain, one of the cyclin-dependent kinase phosphorylation sites and two of the zinc-fingers are not required for the DNA

20   replication activity of ECiz1, when assayed in vitro. It should be noted however that this analysis measures ECiz1 activity in *trans* under conditions where the consequences of mis-localisation are unlikely to be detected. Therefore, it remains possible that the

matrin 3 domain and zinc fingers act in vivo to direct Ciz1 activity to specific sites in the nucleus and thus limit the scope of Ciz1 activity.

Endogenous Ciz1   Antibody V1 recognises Cdc6 as well as p85-Ciz1 (Fig. 1A), so it is
5   not suitable for immuno-fluorescence experiments aimed at visualizing the sub-cellular localization of endogenous Ciz1. We therefore generated two new rabbit polyclonal anti-sera against recombinant ECiz1 fragment Nterm442, designated anti-Ciz1 1793 and 1794. As expected, purified Nterm442 is recognised by anti-Ciz1 antibodies 1793 and 1794 in western blots, but it is also recognised by antibody V1 (Fig. 4A), supporting the
10   conclusion that p85(p100) is indeed Ciz1.

When applied to protein extracts derived from growing 3T3 cells anti-Ciz1 1793 recognised two antigens, with Mr of 125 and 100 kDa (Fig. 4B), whose relative proportions vary from preparation to preparation. The 100 kDa band co-migrates with
15   the cyclin-A responsive antigen that is recognized by antibody V1 (Fig. 1 and 4B), which suggests that both antibodies recognise the same protein in vivo. We confirmed that the p100-Ciz1 bands recognised by antibody V1 and 1793 are the same protein by immuno-precipitation (Fig. 4C). Antibody V1 precipitated a 100 kDa band that was recognised in western blots by 1793, and vice versa. Furthermore, in the same
20   experiment 1793, and to a lesser extent antibody V1, precipitated a 125 kDa antigen, that was recognised in western blots by 1793. Taken together our observations show that the 100 kDa band is indeed Ciz1 (previously known as p85), and they suggest that Ciz1 protein exists in at least two forms in cycling cells.

In addition to the immuno-precipitation evidence described above, several other observations lead to the conclusion that p125 is also a form of Ciz1. First, both of our anti-Ciz1 antibodies (1793 and 1794) have this band in common. Both antibodies

5   produce the same pattern of nuclear staining in immuno-fluorescence experiments, and this is disrupted in cells treated with Ciz1 siRNA (see below). Second, the relative proportions of p100 and p125 vary from preparation to preparation, and could therefore be the result of proteolytic cleavage. Thirdly, our results are strikingly similar to those of Mitsui et al (1999) whose anti-human Ciz1 monoclonal antibody detected two antigens

10   with apparent Mr of 120 and 95 kDa in HEK293 cells. They proposed that the120 kDa form of human Ciz1 protein is processed to produce the 95kDa form and our results are consistent with this proposal.

The 125kDa band recognized by antibody 1793 in mouse and human cells resolves into

15   three Ciz1-related bands during high-resolution electrophoresis of material derived from non-transformed human cells (Wi38-see later), and mouse cells (NIH3T3-not shown). This may be the result of post-translational modification of the Ciz1 protein or of alternative splicing of the Ciz1 transcript.

20   <u>Sub-cellular distribution of Ciz1</u> Anti-Ciz1 1793 was used to visualise the sub-cellular distribution of Ciz1 protein (p85 and p125) in 3T3 cells (Fig. 5A), and in HeLa cells (not shown). In both cell types 1793 reacted with a nuclear-specific antigen, and this was blocked by inclusion of recombinant Nterm442 fragment (Fig. 5B). Unlike Cdc6, which

is shown for comparison (Fig. 5A), Ciz1 is clearly detectable in all 3T3 cells in this

cycling population. Therefore Ciz1 is present in the nucleus throughout interphase,

although minor variations in quantity, or isoform would not be detected by this method.

After detergent treatment overall nuclear Ciz1 staining was reduced in all nuclei, which

5   suggests that Ciz1 is present in the nucleus as both a soluble fraction and also bound to

insoluble nuclear structures.

When soluble protein is washed away, the insoluble, immobilised antigen resolves into a

punctate sub-nuclear speckled pattern at high magnification (Fig. 5C, D). Ciz1 speckles

10   show a similar size range and distribution as replication 'foci' or 'factories', the sites at

which DNA synthesis takes place in S phase. To ask whether Ciz1 is coincident with

sites of replication factories, we compared the position of Ciz1 speckles to the position

of PCNA, a component of replication complexes in S phase cells (Fig. 5C). In confocal

section, PCNA foci are less abundant than Ciz1 foci, but they are almost all co-incident

15   with Ciz1 (Fig. 5D, E, F). This is particularly striking for foci in the medium size range.

In merged images, overlap between the positions of PCNA and Ciz1 foci results in

yellow spots, while the remaining Ciz1 foci that are not co-incident with PCNA are red.

Green (PCNA alone) foci are virtually absent, which suggests that Ciz1 is present at all

sites where DNA replication factories have formed.

20

Ciz1 is also present at sites that don't contain PCNA (Fig. 5D), and unlike PCNA, Ciz1

foci persist throughout interphase (Fig. 5A). One interpretation of these observations is

that Ciz1 marks the positions in the nucleus at which PCNA-containing replication

factories are able to form in S phase, but that not all of these sites are used at the same time. It remains to be determined whether different Ciz1 foci become active sites of DNA replication at different times in S phase, or whether other nuclear activities also occur at sites where Ciz1 is bound. Indeed, at this stage it also remains possible that the

5   100 kDa form and the 125 kDa variants of Ciz1 have different activities, and that they reside at nuclear sites with different functions.

<u>Ciz1 is essential for cell proliferation</u>   So far we have shown that the behaviour of p85 (p100)-Ciz1 correlates with initiation of DNA replication in cell-free assays, that

10   recombinant Ciz1 stimulates the frequency of initiation, and that Ciz1 resides at the same nuclear sites as the DNA replication machinery. However, these data do not show that Ciz1 has an *essential* function in proliferating cells. In order to test this we used RNA interference (RNAi) to selectively reduce Ciz1 transcript levels in NIH3T3 cells. Four target sequences within Ciz1 were chosen (see Fig. 2A) and short interfering (si)

15   RNA molecules were produced in vitro. When applied to cells, all four Ciz1 siRNA's restricted growth (Fig. 6A) and caused a visible reduction in the level of Ciz1 protein after 48 hours (Fig. 6B). The effect of Ciz1 depletion on proliferation becomes apparent between 23 and 40 hours post-transfection, which suggests that the first cell cycle without Ciz1 RNA is relatively unaffected. By 40 hours, controls and Ciz1 siRNA

20   treated cells diverged significantly with no further proliferation in the Ciz1 depleted population. To verify the specificity of Ciz1 depletion, transcript levels were monitored at 24 hours, before proliferation is significantly inhibited (Fig. 6C). At this point Ciz1 transcripts were reduced to 42% of the level in control cells treated with GAPDH

siRNA. These experiments show that Ciz1 is required for cell proliferation and are consistent with a primary function in DNA replication.

To test this further, cells were pulse-labelled with BrdU 48 hours after siRNA treatment

5   to determine the fraction of cells engaged in DNA synthesis (Fig. 6D). When Ciz1 levels were reduced the BrdU labelled fraction was also reduced, suggesting that DNA synthesis is inhibited under these conditions. Furthermore, cells in the Ciz1 depleted population that did incorporate BrdU (approximately 15% of the population) were less intensely labelled. Therefore, in some Ciz1 siRNA treated cells S phase is slowed down

10   rather than inhibited completely, possibly due to incomplete depletion.

Inhibition of DNA synthesis by Ciz1 siRNAs could be a secondary consequence of a general disruption of nuclear function. Therefore, we looked in more detail at a range of other replication proteins whose levels are regulated in a cell cycle dependant manner, to

15   ask whether depleted cells arrest randomly, or accumulate at a particular point.

During initiation of eukaryotic DNA replication Mcm complex proteins assemble at replication origins in late G1, in a Cdc6-dependent manner. Sometime later, DNA polymerases and their accessory factors (including PCNA) become bound to chromatin

20   and origins are activated. This is associated with nuclear export and proteolysis of the majority of Cdc6 and, as DNA synthesis proceeds, gradual displacement of the Mcm complex from chromatin (Bell and Dutta, 2002). In order to identify the point of action of Ciz1 we used immuno-fluorescence to monitor Mcm3 and PCNA. In Ciz1 depleted

44

cells (Fig. 6E, F) both proteins were detectable within the nucleus bound to detergent resistant nuclear structures. Therefore, these factors are unlikely to bind directly to Ciz1, or to be dependent upon Ciz1 for their assembly. In fact, in four independent experiments the average number of cells with detergent-resistant chromatin-bound

5   Mcm3 actually *increased* from 31% (+/-6%) to 51% (+/-5%) (Fig. 6E). Increased Mcm3 indicates that the Ciz1 dependent step occurs after pre-replication complex assembly (but before completion of S phase). In the same cell populations the PCNA positive fraction also increased, from 32% (+/-5%) to 49% (+/-6%) (Fig. 6F), narrowing the point of Ciz1 action to after PCNA assembly. Thus, Ciz1 most likely acts to facilitate

10   DNA replication during a late stage in the initiation process, while failure to act inhibits progression through S phase, leaving Mcm3 and PCNA in place.

Taken together, our cell-free and cell-based investigations paint a consistent picture about the primary function of Ciz1. They suggest that Ciz1 is a novel component of

15   DNA replication factories, and they show that Ciz1 plays a positive role in the mammalian cell-cycle, acting to promote initiation of DNA replication.

Three of our lines of investigation suggest that Ciz1 is required during a late stage in the initiation process after pre-replication complex formation. First, p85 (p100)-Ciz1 antigen

20   accumulates in nuclei exposed to cyclin A-cdk2 concentrations that activate DNA synthesis, implying that Ciz1 functions during this step rather than during earlier replication complex assembly steps (Coverley et al., 2002). Second, functional studies with late G1 nuclei show that recombinant ECiz1 increases the number of nuclei that

incorporate labeled nucleotides in vitro. Therefore, Ciz1 must be active in a step that converts nuclei that are poised to begin DNA synthesis into ones that are actively synthesizing DNA. Third, RNA interference studies point to a Ciz1-dependent step after Mcm complex formation and after PCNA has become assembled onto DNA, but before

5   these proteins are displaced. These distinct lines of investigation lead to strikingly similar conclusions about the point of action of Ciz1 placing it in the later stages of initiation.

Anti-Ciz1 siRNA as a therapeutic strategy   Our analysis shows that Ciz1 is essential for

10   cell proliferation, and that targeting Ciz1 is a viable strategy to restrain proliferation. The alternatively spliced forms of Ciz1 that we observe in various cancers (see below) means that Ciz1 could be targeted in a selective way to restrain proliferation in a subset of cells within a population.

15   By way of example, this could be done by targeting siRNA's to the junction sequence created   in   Ciz1   transcripts   when   the   C-terminal   sequence GTTGAGGAGGAACTCTGCAAGCAG is missing, in small cell lung carcinoma cells, or by using Ciz1 protein lacking the corresponding VEEELCKQ sequence to select specific chemical inhibitors.

20

Accordingly the present invention also provides for the use of junction sequences created in Ciz1 transcripts and proteins when alternatively spliced sequences are not present, as a diagnostic marker, prognostic indicator or therapeutic target.

<u>Embryonic form Ciz1 is localized to the nucleus</u> RT-PCR analysis across potentially variable exons suggest that 3T3 cells predominantly express full-length Ciz1, so our immuno-localization work on endogenous Ciz1 (Fig. 5) does not necessarily reflect the

5   behavior of ECiz1, which lacks several sequence blocks and possibly therefore information that is used to localize the protein. To directly compare the localization of ECiz1 and full-length Ciz1, enhanced GFP tagged constructs were transfected into 3T3 cells (fig. 8A), and microinjected into mouse pro-nuclei (fig. 8B). In all cases tagged Ciz1 and ECiz1 were exclusively nuclear, while a control construct expressing GFP

10 alone was present in the nucleus and the cytoplasm. GFP-Ciz1 and GFP-ECiz1 were both visible in live cells as sub-nuclear foci, similar to replication foci seen in fixed cells by immuno-fluorescence. Thus, the three sequence blocks that are absent from ECiz1 do not appear to contribute to the nuclear localization of Ciz1.

15 Over the three day period following transfection no cell division was observed in the GFP-Ciz1 and GFP-ECiz1 transfected cells. These data suggest that overexpression of functional Ciz1 has an inhibitory effect on the cell cycle (in cells that have their regulatory pathways intact).

20 <u>Coalescence</u> When GFP-tagged constructs in which the C-terminal one third of Ciz1 had been removed were transfected into 3T3 cells, differences between ECiz1 and full length Ciz1 were observed (fig. 8C). By 48 hours FL Ciz1 N-term(442 equivalent) had coalesced into large intra-nuclear blobs which only became apparent in the ECiz1 N-

term442 transfected population by day 3 or later. Before this time ECiz1 N-term442 was localised as a nuclear specific but diffuse pattern. Thus ability to coalesce is quantifiably different between Ciz1 and ECiz1, and is therefore affected by one of the three alternatively spliced exons (2/3, 6 or 8).

5

Like cells transfected with full length Ciz1 and ECiz1, cells transfected with constructs in which the C terminal one third was removed were not seen to multiply during the three day monitoring period.

10 <u>C-terminal domains anchor Ciz1 to nuclear structures</u> As described above, the difference between Ciz1 and ECiz1 N-term is masked when C-terminal domains are also present (fig. 8A). Furthermore the C-terminal fragment alone directs GFP tag to chromatin, forming an irregular pattern that is not as spotty (focal) as Ciz1 or ECiz1, but which remains attached to chromosomes during mitosis (fig. 8D). This suggests that C-

15 terminal domains are involved in immobilizing Ciz1 on a structural framework in the nucleus. Notably, cells transiently transfected with C-terminal fragment continued to divide resulting in gradual dilution of green fluorescence.

<u>Ectopic Ciz1 promotes premature entry to S phase</u> We looked at events occurring during

20 the first day after transfection. The S phase fraction in transfected cells (green) was compared to the S phase fraction in untransfected cells, by labelling with BrdU at various intervals. During long labelling windows including 0-22 hours (fig. 8E), 0-12 hours and 0-7 hours (not shown), consistently more of the Ciz1 and ECiz1 transfected

cells were engaged in DNA synthesis, compared to untransfected cells. This suggests that Ciz1 and ECiz1 have a positive effect on the G1-S transition, promoting unscheduled entry to S phase. Similar results were obtained with 3T3 cell populations that were densely plated before transfection. This was done in order to minimise the

5 fraction in the untransfected population that was engaged in S phase as part of the normal cell cycle. Under these conditions the difference between the transfected and untransfected population was maximised, clearly demonstrating the effect of ectopic Ciz1 on initiation of DNA replication.

10 Conversely, when cells were labelled with BrdU during a short pulse administered at 22 hours (fig. 8E), or at 10 hours or 12 hours post-transfection (not shown), the labelled fraction was consistently *reduced* in the Ciz1 and ECiz1 transfected populations. This suggests that the S phase that is induced by ectopic Ciz1 or ECiz1 is abnormal, with slow or aborted DNA synthesis that is not sufficient to label cells during short windows

15 of exposure to BrdU.

Therefore, ectopic Ciz1 and ECiz1 have two effects on S phase in cultured cells. They promote DNA replication, but this results in slow or aborted DNA synthesis.

20 <u>Clones with altered proliferation potential</u> We also monitored transfected populations of 3T3 cells over a three week time period. In cells transfected with the GFP-Nterm442 or the non-alternatively spliced equivalent and maintained under selection with G418, large foci containing hundreds of cells were observed (fig. 9A). These clusters contained

large numbers of GFP expressing cells, demonstrating that over-expression of the N-terminal portion of ECiz1 (in which replication activity resides) is not lethal, and suggesting that over-expression leads to altered proliferation phenotype, compared to untransfected cells, including loss of contact inhibition and failure to form a monolayer.

5   This Ciz1-dependent altered behaviour could contribute to tumour formation. A similar truncated version of mouse Ciz1, lacking putative chromatin interaction domains was previously isolated from a mouse melanoma (fig. 2).

Human Ciz1 and cancer

10   Ciz1 cDNAs in public databases   As mentioned above human Ciz1 is alternatively spliced at the RNA level to yield transcripts that lack three of the same exons as mouse embryonic Ciz1. Seven human Ciz1 cDNAs have been recorded in public databases (fig. 10), submitted by Mitsui et al (1999), Warder and Keherly (2003) and large-scale genome analysis projects (NIH-MGC project, NEDO human cDNA sequencing project). Only one is derived from normal adult tissue, and this contains all predicted exons

15   (AB030835). The rest are derived from embryonic cells (AK027287), or notably from four different types of paediatric cancer (medulloblastoma, AF159025, AF0234161, retinoblastoma, AK023978, neuroblastoma, BC004119 and burkitt lymphoma, BC021163). The embryonic form and the cancer derived forms lack sequence blocks

20   from the same three regions as our embryonic mouse clone, and from a fourth region which corresponds to exon 4. Therefore, the limited data suggests that alternatively spliced forms are more prevalent early in development. This correlation has not previously been noted in the scientific literature. The presence of alternatively spliced

Ciz1 in paediatric cancers raises the possibility that Ciz1 mis-splicing might be linked to inappropriate cell proliferation.

For example, one of the variable exons encodes a short conserved DSSSQ sequence motif that is absent in mouse ECiz1 and in a human medulloblastoma. This is directly adjacent to the consensus cdk phosphorylation site that we have shown to be involved in regulation of ECiz1 function. Conditional inclusion of the DSSSQ sequence might make Ciz1 the subject of regulation by the ATM/ATR family of protein kinases, which phosphorylate proteins at SQ sequences, thereby restraining Ciz1 initiation function in response to DNA damage.

Analysis of expressed sequence tags. The presence of alternatively spliced Ciz1 in paediatric cancers prompted a detailed analysis of Ciz1 ESTs. There are 567 expressed sequence tags (ESTs) included in NCBI unigene cluster Hs.23476 (human Ciz1). These are derived from a wide range of normal and diseased tissues and cell lines. Sequences have been translated and mapped against the predicted full-length amino-acid sequence of human Ciz1. Sequence alterations that give rise to amino-acid substitutions, deletions, frame-shifts and premature termination of translation have been recorded.

Alternatively spliced Ciz1 variants were also seen in this EST data set and are recorded here. The four sequence blocks that we previously reported to be alternatively spliced in human and mouse Ciz1 (Fig. 2) were observed in the EST sequences, as well as a previously undetected variant that lacks the exon 14 derived sequence VEEELCKQ. All

of these recurrently variant sequence blocks are bounded by appropriate splice sites. A

sixth variable sequence block was identified in one carcinoma derived library, caused by

inclusion   of

GCCACCCACACCACGAAGAGATGTGTTTGCCCACGTTCCAGTGCAGGGGTG

5   GAGCACAGCCCGGCTTGTTACAGATAT.

ESTs are grouped according to the cell type from which they were derived with the

primary divisions occurring between neoplastic cells of adult, childhood or embryonic

origin. ESTs from normal tissue of embryonic or adult origin are included for

10   comparison.   EST-derived Ciz1 protein maps are shown in fig. 11A-E and the

alternatively spliced exons summarized in fig. 11F.

Three sequence blocks in the N-terminal end of human Ciz1 are absent in transcripts

from medulloblastomas and neuroblastoma (fig. 11A), and occasionally absent from

15   Ciz1 transcripts from other cancers. We also found similar alternative splicing in a third

paediatric cancer, Ewings sarcoma (see below). Paediatric cancer-associated

alternatively spliced sequences are from exons 2/3 (at least two versions), exon 4 and

exon 6.

20   Exon 8 variants in which one or more copies of a Q-rich degenerate repeat are absent

have been noted in transcripts derived from normal cells (of embryonic or adult neural

origin) and from various cancers. Alternative splicing in this region could produce Ciz1

with inappropriate activity, therefore exon 8 variant expression, or occurance of point

mutations which influence splicing in this region, might be useful as diagnostic or prognostic markers in cancer. The alternatively spliced degenerate repeats in exon 8 are detailed below and summarised in fig. 11F.

5 In the C-terminal half of the human Ciz1 protein two sequence blocks are variably spliced. One of these is missing from transcripts derived from three out of five lung carcinoma and lung carcinoid libraries, and from three other carcinoma libraries (but very rarely from transcripts from other cell types).

10 The second variant sequence block is due to inappropriate inclusion of extra sequence in transcripts from the epidermoid carcinoma library (MGC102).

These sequences and the junction sequences formed in Ciz1 proteins, and Ciz1 transcripts when these segments are excluded or included, are potential targets for 15 selective inhibition of cell proliferation in a wide range of different cancers. The remaining non-variant sequences are potential targets for non-selective inhibition of cell proliferation.

In addition to splicing variations, other non-typical Ciz1 transcripts were found to 20 preferentially occur in some cancers. In Rhabdomyosarcomas Ciz1 is prematurely terminated leading to a predicted protein that lacks C-terminal nuclear binding domains. This could lead to inappropriate DNA replication and might therefore be a therapeutic target or marker in this type of cancer.

Several transcripts contain point mutations that lead to amino-acid substitutions in putative cyclin-dependent kinase (cdk) phosphorylation sites. In the cervical carcinoma library MGC12, this occurs twice. We have shown that two cdk phosphorylation sites

5   are involved in restraining Ciz1 activity (fig. 3C and D), implicating these mutations in the deregulation of proliferation in cancer cells. One of these is the same as the carcinoma-derived mutant mentioned above (fig. 11E). Cancer-derived transcripts with point mutations in Ciz1 could also be targeted by RNA interference, or have value as diagnostic or prognostic indicators.

10

Investigation of Ciz1 variant expression in paediatric cancers

Ciz1 variant expression was investigated in 6 Ewings Sarcoma family tumour cell lines (ESFTs) and two neuroblastoma cell lines, using RTPCR with primer sets that span three regions of known Ciz1 variability (fig. 12A). This analysis showed that the pattern

15   of Ciz1 variant expression is different in ESFT cells compared to neuroblastoma cells compared to non-transformed cells, but apparently very similar within sets of cell lines from the same tumour. Therefore, Ciz1 variant expression could have prognostic or diagnostic potential for these cancers. Minor variations within a set of lines from the same tumour type could have prognostic value.

20

By subcloning and sequencing amplified transcripts we found that all six ESFT lines tested express an exon 4 minus form of Ciz1. As Ciz1 is essential for cell proliferation (see below), this offers a possible route for selective restraint of ESFT cells. Transcripts

from the two neuroblastoma cell lines tested rarely lack exon 4 but frequently lack sequences the DSSSQ motif encoded by exon 6 (fig. 12B).

This experimental analysis confirms that paediatric cancers express forms of Ciz1 with
5 variable inclusion of exons 4, 6 and probably exons 2/3.

Two versions of the sequence encompassing exon 8 and one form of the sequence encompassing the VEEELCKQ-coding sequence were detected in ESFTs, neuroblastomas and control suggesting that these regions do not contribute to
10 deregulation of Ciz1 in these paediatric cancers.

In all cases, Ciz1 RT-PCR products were most abundant in reactions carried out with RNA samples from cancer cell lines, compared to controls (Wi38, HEK293, NIH3T3 cells, and primary human osteoblasts). This is consistent with increased expression of
15 Ciz1 variants in tumours.

Analysis of Ciz1 protein expression in prostate cancer cell lines

Normal, non-transformed human lung fibroblasts (and mouse NIH3T3 cells) express two major forms of Ciz1 that are detected by anti-Ciz1 polyclonal antibody 1793 in western
20 blots (fig. 13A). The larger (approximately 125kDa) band resolves into three distinct bands that are present in equal proportions in Wi38 cells, but grossly uneven proportions in prostate cancer cell lines PC3 and LNCAP (and ESFT cell lines –not shown). We postulate that these protein isoforms are generated by expression of variably spliced

exons. Both tumour cell lines also contain more Ciz1 antigen than Wi38 cells, consistent with over-expression of Ciz1 in these cancer cell lines.

5   Taken together our results (experimental and bioinformatics analysis of genome data) support the conclusion that Ciz1 is mis-regulated in a wide range of human cancers. We have shown that the Ciz1 protein plays a positive role in the DNA replication process, therefore mutant Ciz1 could contribute to cellular transformation, rather than be a consequence of it. If deregulation of Ciz1 is a common step in this process it represents a very attractive target for development of therapeutic agents.

10

We have also associated particular changes with specific cancers, making it a real possibility that Ciz1 could be useful as a diagnostic or prognostic marker.

These include :-

- Alternative splicing in the N-terminal part of the protein (that contains

15   replication activity in vitro) in paediatric cancers.

- Point mutations in cyclin-dependent kinase phosphorylation sites known to be involved in restraining Ciz1 replication activity.

- Non-typical expression and nuclear binding properties of Ciz1-p125 forms in prostate carcinoma cell lines, possibly due to mis-regulated splicing of the

20   degenerate repeats in exon 8, or other exons.

- Conditional exclusion of a discrete motif (VEEELCKQ) in the C-terminal end of Ciz1 (probably involved in localization of Ciz1 protein within the nucleus) in small cell carcinoma of the lung and other carcinomas.

- Increased levels of Ciz1 protein and RNA (detected by Western blot and by RT-PCR) in all cancer derived cells lines tested so far, compared to Wi38 normal embryonic lung fibroblast, human osteoblast RNA and mouse NIH3T3 fibroblasts.

5

The sequences shown in figures 14 to 21 are of use for the development of therapeutic,

10   diagnostic, or prognostic reagents.

Materials and Methods

Cloning. A lamba triplEx 5'-stretch, full length enriched cDNA expression library derived from 11 Day old mouse embryos (Clonetech ML5015t) was used to infect *E.*

15   *coli* Xl1blue according to the recommended protocol (Clonetech). Plaques were lifted onto 0.45 micron nitrocellulose filters pre-soaked in 10mM IPTG (Sigma). Affinity purified antibody V1 was applied to approximately 3×10^6 plaques at 1/1000 dilution in PBS, 10% non-fat milk powder, 0.4% Tween20, after blocking for 30 minutes in the absence of antibody. After two hours filters were washed three times with the same

20   buffer and reactive plaques were visualized with anti-rabbit secondary antibody conjugated to horse-radish peroxidase (Sigma), and enhanced chemi-luminescence (ECL, Amersham) according to standard procedures. 43 independent plaques were

picked but only two strains of phage survived a further three rounds of screening. These were converted to pTriplEx by transforming into BM25.8 and sequenced. One codes for mouse Cdc6 (clone P) and the other (clone L) for an unknown mouse protein that is homologous to human Ciz1. We refer to this as embryonic Ciz1 (ECiz1) and it was

5   submitted to EMBL under the accession number AJ575057.

Bacterial expression pGEX based bacterial expression constructs (Amersham) were used to produce ECiz1 proteins for in vitro analysis. pGEX-ECiz1 was generated by inserting a 2.3kb SmaI-XbaI (blunt ended) fragment from clone L into the SmaI site of pGEX-6P-

10   3. pGEX-Nterm442 was generated by inserting the 1.35kb XmaI-XhoI fragment into XmaI-XhoI digested pGEX-6P-3, and pGEX-Cterm274 by inserting the 0.95kb XhoI fragment into XhoI digested pGEX-6P-3. pGEX-T(191/2)A was generated from pGEX-ECiz1 by site directed mutagenesis (Stratagene Quikchange) using primers AACCCCCTCTTCCGCCGCCCCAATCGCAAGA   and

15   TCTTGCGATTGGGGGCGGCGGAAGAGGGGGTT. pGEX-T(293)A was generated from pGEX-ECiz1 using primers AAGCAGACACAGGCCCCGGATCGGCTGCCT and AGGCAGCCGATCCGGGGCCTGTGTCTGCTT. Integrity and reading frame of all clones were sequence verified.

20   Recombinant Ciz1, Ciz1 fragments and point mutants were produced in BL21-pLysS (Stratagene) as glutathione S-transferase-tagged protein. This was purified from sonicated and cleared bacterial lysates by binding to glutathione sepharose 4B (Amersham). Recombinant protein was eluted by cleavage from the GST tag using

precision protease (as recommended by the manufacturer, Amersham), into buffer

(50mM Tris-HC pH 7.0, 150mM NaCl, 1mM DTT). This yielded protein preparations

between 0.2 and 2.0 mg/ml. For replication assays serial dilutions were made in 100mM

Hepes pH 7.8, 1mM DTT, 50% glycerol so that not more than1ml of protein solution

5   was added to 10ml replication assays, yielding the concentrations shown. Consistent

with previous observations (Mitsui et al., 1999; Warder and Keherly, 2003) recombinant

Ciz1, and derived fragment N-term442 migrated through SDS-PAGE with anomalously

high molecular weight. Cyclin A-cdk2 was produced in bacteria as previously described

(Coverley et al., 2002).

10

Anti-Ciz1 antibodies Rabbit polyclonal antibody V1 (Coverley et al., 2000; Stoeber et

al., 1998; Williams et al., 1998) was raised against an internal fragment of bacterially

expressed human Cdc6 corresponding to amino-acids 145-360, and affinity purified by

standard procedures (Harlow and Lane, 1988). This antibody reacts strongly with

15   endogenous p100-Ciz1 and also with ECiz1 Nterm442 fragment. Alignment of

Nterm442 with Cdc6 amino-acids 145-360 suggest that the shared epitope could be at

294-298 or 304-312 in mouse Ciz1. Recombinant Nterm442 was used to generate two

Ciz1-specific polyclonal anti-sera designated 1793 and 1794 (Abcam). 1793 has been

used routinely in the experiments described here. Its specificity was verified by

20   reciprocal immuno-precipitation and western blot analysis with antibody V, by inclusion

of Nterm 442 (25µg/ml in antibody buffer, 10mg/ml BSA, 0.02% SDS, 0.1% Triton

X100 in PBS), which blocked reactivity with endogenous epitopes, and by siRNA-mediated depletion of Ciz1 that specifically reduced 1793 nuclear staining.

<u>Immunoprecipitation</u> Asynchronousy growing 3T3 cells were washed in PBS, rinsed in

5   extraction buffer (20mM Hepes pH7.8, 5mM potassium acetate, 0.5mM magnesium

chloride) supplemented with EDTA-free protease inhibitor cocktail (Roche) and scrape

harvested as for replication extracts. Cells were lysed with 0.1% Triton X 100 and the

detergent resistant pellet fraction extracted with 0.3M NaCl in extraction buffer. 5µl of

1793 or 2µl of antibody V were used per 100µl of extract and incubated for 1 hour at

10   4°C. Antigen-antibody complexes were extracted with 100µl of protein G-sepharose

(Sigma) and beads were washed five times with 50mM Tris pH 7.8, 1mM EDTA, 0.1%

NP40, 150mM NaCl. Complexes were boiled in loading buffer (100mM DTT, 2% SDS,

60mM Tris pH6.8, 0.001% bromophenol blue) and resolved by 6.5% SDS-polyacrylamide gel electrophoresis.

15

<u>Immuno-fluorescence</u> Cells were grown on coverslips and fixed in 4%

paraformaldehyde, with or without brief pre-exposure to 0.05% Triton X100 in PBS.

Endogenous Ciz1 was detected with 1793 serum diluted 1/2000 in antibody buffer

following standard procedures. Mcm3 was detected with monoclonal antibody sc9850

20   (1/1000), Cdc6 with monoclonal sc9964 (1/100) and PCNA with monoclonal antibody

PC10 (1/100, all Santa Cruz Biotechnology). Co-localisation analysis of dual stained

fluorescent confocal images was carried out as described (Rubbi and Milner, 2000; van

Steensel et al., 1996).

Cell synchrony Mouse 3T3 cells were synchronized by release from quiescence as

5   previously described (Coverley et al., 2002). Nuclei prepared from cells harvested 17

hours after release (referred to as 'late-G1') were used in all cell-free replication

experiments described here. This yielded populations containing S phase nuclei,

replication competent late G1 nuclei and unresponsive early G1/G0 nuclei, in varying

proportions. Recipient, mid-G1 3T3 extracts were prepared at 15 hours (these typically

10   contain approximately 5% S phase cells). The series of cell-free replication experiments

described here required large amounts of standardized extract, therefore HeLa cells were

used because they are easily synchronized in bulk. S phase HeLa extracts were prepared

from cells released for two hours from two sequential thymidine-induced S phase

blocks, as described (Krude et al., 1997).

15

Cell-free DNA replication   DNA replication assays were performed as described

(Coverley et al., 2002; Krude et al., 1997). Briefly, 10μl of mid G1 or S phase extract

(supplemented with energy regenerating system, nucleotides and biotinylated dUTP),

and 5×10^4 late G1 phase nuclei were incubated for 60mins at 37°C. Reactions were

20   supplemented with baculovirus lysate containing cyclin A-cdk2 (Fig.1 B and C), where

0.1 μl of lysate has the same specific activity as 1nM purified kinase (Coverley et al.,

2002). All recombinant proteins were serially diluted in 100mM Hepes pH 7.8, 1mM

DTT, 50% glycerol, so that not more than 1μl was added to 10μl replication assays, generating the concentrations indicated. Reactions were stopped with 50μl of 0.5% Triton X100 and fixed by the addition of 50 μl of 8% paraformaldehyde, for 5 minutes. After transfer to coverslips nuclei were stained with streptavidin-FITC (Amersham) and

5   counterstained with Toto-3-iodide (Molecular Probes). The proportion of labelled nuclei was quantified by inspection at 1000X magnification, and all nuclei with fluorescent foci or intense uniform labelling were scored positive. Images of *in vitro* replicating nuclei were generated by confocal microscopy at 600X magnifications, of samples counterstained with propidium iodide. For analysis of nuclear proteins, nuclei were re-

10   isolated after 15 minutes exposure to initiating conditions, by diluting reactions two fold with cold PBS and gentle centrifugation.

Data analysis and presentation Prior to use in initiation assays each preparation of synchronized G1 phase nuclei is tested so that the proportion of nuclei that are already in

15   S phase is established ('%S'). To do this nuclei are incubated in an extract that is incapable of inducing initiation of DNA synthesis (from mid-G1 phase cells harvested 15 hours after release from quiescence), but that will efficiently support elongation DNA synthesis from origins that were initiated *in vivo*. The elongating fraction of nuclei incorporates labeled nucleotides efficiently during *in vitro* initiation assays but is

20   uninformative. Routinely this fraction is pre-established and subtracted from the raw data. Synchronized populations in which 20% or less are in S phase are used for initiation assays.

When 3T3 cells are released from quiescence by the protocol used here no more than 70% of the total population enters S phase (Coverley et al., 2002). However, the highest observed replication frequency in vitro is nearer 50%; usually obtained by incubation

5   with ECiz1. For the G1 population of 3T3 nuclei used here 17% were in S phase (%S) and the maximum number that replicated in any assay in vitro was 51% (% replication). Therefore, 34% of this population is *competent* to initiate replication in vitro (%C). Thus, for each data point in Figs. 3B-F, % initiation = (% replication -%S)/%C x 100.

10   <u>RNA interference</u> Endogenous Ciz1 was targeted in proliferating NIH3T3 cells using in vitro transcribed siRNAs (Ambion Silencer kit), directed against four regions of mouse Ciz1. Oligonucleotide sequences that were used to generate siRNAs are AAGCACAGTCACAGGAGCAGACCTGT CTC and AATCTGCTCCTGTGACTGTGCCCTGTCTC for siRNA 4, AATCTGTCAC

15   AAGTTCTACGACCTGTCTC and AATCGTAGAACTTGTGACAGACCTGTCTC for siRNA 8, AATCGCAAGG ATTCTTCTTCTCCTGTCTC and AAAGAAGAAGAA TCCTTGCGACCTGTCTC for siRNA 9, and AATCTGCAGCAGTTCTTTCCCCCTGTCTC and AAGGGAAAGAACTGCTGCAGACCTGTCTC for siRNA 11. Target sequences that

20   are distributed throughout the Ciz1 transcript were chosen based on low secondary structure predictions and on location within exons that are consistently expressed in all known forms of Ciz1 (sequences 4, 8, 11), with the exception of one (siRNA 9) that is known to be alternatively spliced. Negative controls were untreated, mock treated

(transfection reagents but no siRNA) and cells treated with GAPDH siRNA (Ambion). Cy3 labelled siRNAs (Ambion) were used to estimate transfection efficiency, which was found to be greater than 95%. RNA interference experiments were performed in 24 well format starting with $2x10^4$ cells per well in 500μl of medium (DMEM with glutamax

5   supplemented with 4% FCS). siRNA's were added 12 hours after plating using oligofectamine reagent for delivery (Invitrogen). Unless stated otherwise, siRNAs were used in pairs (at 2nM total concentration in medium), as two doses with the second dose delivered in fresh medium 24 hours after the first. Results were assessed at 48 hours after first exposure, by counting cell number, S phase labelling, and immuno-staining.

10   Northern blots were performed on RNAs isolated from cells treated for 24 hours with a single dose of siRNA, in reactions that were scaled up 5 fold. RNA was prepared using Trizol Reagent (Invitrogen) and samples were electrophoresed through 1% agarose, transferred onto Hybond N+ nylon membrane (Amersham), and sequentially hybridised at 50°C with cDNA probes using NorthernMax kit reagents (Ambion), following

15   manufacturers instructions. The membrane was stripped between each hybridisation using 0.5% SDS solution at 90°C, allowed to cool slowly to room temperature. Probes were [^{32}P]-dCTP labelled using Random Primers DNA labelling system (Gibco BRL), and used in the following order: i. A 1.35kb Xma1-Xho1 fragment derived from ECiz1. ii. Human β-actin cDNA (Clontech) and iii. Mouse GAPDH cDNA (RNWAY

20   laboratories). The membrane was washed twice in 2X SSC 0.2% SDS for 30-60 mins each, followed by one wash in 0.2X SSC 0.2% SDS for 30 mins, at 55-65°C, depending on probe used. Hybridisation signals were quantified using an Amersham Biosciences

Typhoon 9410 variable mode imager, and Image Quant TL software (v2002). Band

intensities are expressed in arbitrary units (in parentheses), and results for Ciz1 and

GAPDH were normalised against those for β-actin, and expressed as a %.

5   S phase labelling The fraction of nuclei undergoing DNA synthesis in vivo was

monitored by supplementing culture medium with 20 μM bromodeoxyuridine (BrdU,

Sigma) for 20 minutes. Incorporated BrdU was visualized after acid treatment with

FITC-conjugated anti-BrdU monoclonal antibody (Alexis Biochemicals) according to

manufacturers instructions. Nuclei were counterstained with Hoescht 33258 and scored

10   under high (1000X) magnification.

Green fluorescent protein tagged Ciz1

Full-length mouse Ciz1 cDNA was obtained from UK HGMP Resource Centre (MGC

clone 27988) and the sequence fully verified. A 2.8kb SmaI-XbaI (blunt ended) full

15   length Ciz1 fragment from this clone, and a 2.3kb SmaI-XbaI (blunt ended) ECiz1

fragment from pTriplEx-clone L were ligated in frame with enhanced green fluorescent

protein (EGFP) into the SmaI site of pEGFP-C3 (Clontech). pEGFP-C3 with no insert

was used as a control. Constructs were transfected into NIH3T3 cells using *Trans*IT-293

(Mirus), following manufacturers instructions or microinjected into the male pro-nucleus

20   of fertilized mouse eggs at the one cell stage. Growing 3T3 cells transfected with full

length EGFP-Ciz1, or EGFP-ECiz1 were analysed by live cell fluorescent microscopy

up to three days after transfection. DNA synthesis was monitored during the first 24

hours after transfection, by including the nucleotide analogue BrdU in cell culture medium for various time periods as indicated in figure legends. As described above any cells undergoing DNA synthesis while exposed to BrdU stain with anti- BrdU monoclonal antibody generating red nuclei.

5   Ciz1 transfected cells were also maintained under selection with 50 μg/ml G418, in standard culture medium (DMEM Glutamax plus 10% fetal calf serum) for up to a month, yielding cell populations with altered morphology.

EST sequence analysis

10   Individual expressed sequence tags (ESTs) mapping to NCBI unigene cluster Hs.23476 (human Ciz1) were translated using Genejockey and the predicted amino-acid sequence compared to the predicted sequence for full length Ciz1, with the aim of identifying recurrent changes in cancer cells. In order to exclude errors that reflect poor quality DNA sequence such as that which occurs at the end of long sequencing runs, only those

15   changes positioned more than 8 amino-acids from the end of uninterrupted sequence are included in this analysis. Frame-shifts that are restored by a second alteration later in the read, and frame-shifts that are followed by a stop codon are only included if followed by uninterrupted sequence. Thus the majority of sequencing errors are excluded from this analysis. However, it is expected that many of the point mutations that remain (including

20   frame-shifts and stops) reflect errors introduced during sequencing. Therefore, this analysis is aimed at uncovering trends, with weight being given to point mutations only if they appear more than once.

Of 567 sequences, that map to Ciz1 unigene cluster we have analysed most (all paediatric cancers, prostate and lung carcinomas, leukemias and lymphomas and a wide range of non-diseased tissues). Some were not mapped because they are extremely short

5   reads or yielded very short amino-acid sequences upon translation, and for a small number we detected no homology to the Ciz1 coding sequence. A small number of ESTs were excluded from the analysis because of multiple frameshifts that produced stretches of homology in all three frames, with no indication of the reading frame used in vivo. These were all from cancer derived material, usually adenocarcinomas.

10

RT-PCR analysis of Ciz1 isoform expression RNA was isolated using trizol reagent following recommended procedures, DNAse treated and reverse transcribed using random hexamers and superscript II, then amplified with Ciz1 specific primers: -
h/m5 CAGTCCCCACCACAGGCC,

15   h/m2 GGCTTCCTCAGACCCCTCTG.

H/m3 ACACAGACCTCTCCAGAGCACTTAG

H/m4 ATGGTGACCTTCAGGGAGC

H4 TCCTTGGCGA TGTCCTCTGG GCAGG

H3 TCCCTCCTCA ACGGCTCCAT GCTGC

20   H6 CG TGGGGGCGAC TTGAGCGTTG AGG

H1 GATGCCAGGGGT ATGGGGCGCC GGG

H2 TCCGAGCCCT TCCACTCCTC TCTGG

Analysis of Ciz1 protein isoforms in cancer cell lines

Cells were grown in DMEM with 10% FCS until sub-confluent, rinsed in cold hepes buffered saline supplemented with EDTA free protease inhibitor cocktail (Roche) then scrape harvested and supplemented with 0.1% Triton X100. Detergent-insoluble

5 material (including nuclei) was pelleted by gentle centrifugation to yield supernatant (SN) and pellet fractions (P). These were boiled in reducing SDS-PAGE sample buffer and proteins resolved by electrophoresis through 8% SDS-PAGE. After transfer to nitrocellulose, Ciz1 isoforms were detected with anti-Ciz1 antibody 1793). All methods used in this analysis are well documented elsewhere.

10

References

Bell, S. P. and Dutta, A. (2002). DNA replication in eukaryotic cells. *Annu Rev Biochem* **71**, 333-74.

Cook, P. R. (1999). The organization of replication and transcription. *Science* **284**,

15 1790-5.

Corpet, F. (1998). Multiple sequence alignment with hierarchical clustering. *Nucl. Acids Res.* **16**, 10881-10890.

Coverley, D., Laman, H. and Laskey, R. A. (2002). Distinct roles for cyclins E and A during DNA replication complex assembly and activation. *Nat Cell Biol* **4**, 523-8.

20 **Coverley, D., Pelizon, C., Trewick, S. and Laskey, R. A.** (2000). Chromatin bound Cdc6 persists in S and G2 phases in human cells, while soluble Cdc6 is destroyed in a cyclin A-cdk2 dependent process. *J. Cell Sci.* **113**, 1929-1938.

Fujita, M. (1999). Cell cycle regulation of DNA replication initiation proteins in mammalian cells. *Front Biosci* 4, D816-23.

Hanahan, D. and Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell* 100, 57-70.

Harlow, E. and Lane, D. (1988). Antibodies : A laboratory manual. New York: Cold

5   Spring Harbour Laboratory Press.

Jones, D. L., Alani, R. M. and Munger, K. (1997). The human papillomavirus E7 oncoprotein can uncouple cellular differentiation and proliferation in human keratinocytes by abrogating p21Cip1-mediated inhibition of cdk2. *Genes Dev.* 11, 2101-2111.

10   Krude, T. (2000). Initiation of human DNA replication *in vitro* using nuclei from cells arrested at an initiation-competent state. *J. Biol. Chem.* 275, 13699-13707.

Krude, T., Jackman, M., Pines, J. and Laskey, R. A. (1997). Cyclin/Cdk-dependent initiation of DNA replication in a human cell-free system. *Cell* 88, 109-119.

Laman, H., Coverley, D., Krude, T. K., Laskey, R. A. and Jones, N. (2001). Viral

15   cyclin/cdk6 complexes initiate nuclear DNA replication. *Mol. Cell. Biol.* 2, 624-635.

Mercatante, D. R. and Kole, R. (2002). Control of alternative splicing by antisense oligonucleotides as a potential chemotherapy: effects on gene expression. *Biochim Biophys Acta* 1587, 126-32.

Mitsui, K., Matsumoto, A., Ohtsuka, S., Ohtsubo, M. and Yoshimura, A. (1999).

20   Cloning and characterization of a novel p21cip1/wafl-interacting zinc finger protein, Ciz1. *Biochem. Biophys. Res. Com.* 264, 457-464.

Nakayasu, H. and Berezney, R. (1991). Nuclear matrins: identification of the major nuclear matrix proteins. *Proc Natl Acad Sci U S A* 88, 10312-6.

Ohnuma, S., Philpott, A. and Harris, W. A. (2001). Cell cycle and cell fate in the nervous system. *Curr Opin Neurobiol* 11, 66-73.

Parker, S. B., Eichele, G., Zhang, P., Rawls, A., Sands, A. T., Bradley, A., Olson, E. N., Harper, J. W. and Elledge, S. J. (1995). p53-independent expression of p21Cip1 in
5   muscle and other terminally differentiating cells. *Science* 267, 1024-7.

Rubbi, C. P. and Milner, J. (2000). Non-activated p53 co-localizes with sites of transcription within both the nucleoplasm and the nucleolus. *Oncogene* 19, 85-96.

Sherr, C. J. and Roberts, J. M. (1999). CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev.* 13, 1501-1512.

10   Stoeber, K., Mills, A. D., Kubota, Y., Krude, T., Romanowski, P., Marheineke, K., Laskey, R. A. and Williams, G. H. (1998). Cdc6 protein causes premature entry into S phase in a mammalian cell-free system. *EMBO J.* 17, 7219-7229.

van Steensel, B., van Binnendijk, E. P., Hornsby, C. D., van der Voort, H. T., Krozowski, Z. S., de Kloet, E. R. and van Driel, R. (1996). Partial colocalization of
15   glucocorticoid and mineralocorticoid receptors in discrete compartments in nuclei of rat hippocampus neurons. *J Cell Sci* 109 (Pt 4), 787-92.

Warder, D. E. and Keherly, M. J. (2003). Ciz1, Cip1 interacting zinc finger protein 1 binds the consensus DNA sequence ARYSR(0-2)YYAC. *J Biomed Sci* 10, 406-17.

Williams, G. H., Romanowski, P., Morris, L., Madine, M., Mills, A. D., Stoeber, K.,
20   Marr, J., Laskey, R. A. and Coleman, N. (1998). Improved cervical smear assessment using antibodies against proteins that regulate DNA replication. *Proc. Natl. Acad. Sci. USA* 95, 14932-14937.

Zezula, J., Casaccia-Bonnefil, P., Ezhevsky, S. A., Osterhout, D. J., Levine, J. M., Dowdy, S. F., Chao, M. V. and Koff, A. (2001). p21cip1 is required for the differentiation of oligodendrocytes independently of cell cycle withdrawal. *EMBO Rep* 2, 27-34.

5

10

15